The Crop Journal xxx (xxxx) xxx

Contents lists available at ScienceDirect

The Crop Journal

journal homepage: www.elsevier.com/locate/cj

Identification of loci and candidate genes controlling kernel weight in barley based on a population for which whole genome assemblies are available for both parents

Hong Zhou [a,b,1], Wei Luo [a,b,1], Shang Gao [b,c,1], Jian Ma [a], Meixue Zhou [c], Yuming Wei [a], Youliang Zheng [a], Yaxi Liu [a,*], Chunji Liu [b,*]

[a] Triticeae Research Institute, Sichuan Agricultural University, Wenjiang, Chengdu 611130, Sichuan, China
[b] CSIRO Agriculture Flagship, 306 Carmody Road, St. Lucia, QLD 4067, Australia
[c] School of Land and Food and Tasmanian Institute of Agriculture, University of Tasmania, Hobart 7250, Australia

ARTICLE INFO

Article history:
Received 5 May 2020
Revised 1 July 2020
Accepted 10 September 2020
Available online xxxx

Keywords:
Barley
Kernel weight
Near isogenic lines
AWCS276

ABSTRACT

Kernel weight (KW), together with kernel number per unit area, determines yield of cereal crops. Here, two barley recombinant inbred lines (RILs) populations with a shared parent were used to identify loci controlling KW. One is Baudin/AWCS276 (BA) for which a linkage map was available. Several large-effect QTL controlling KW were detected in this population. Another is Morex/AWCS276 (MA). A linkage map with 5273 makers formed 1454 clusters, was constructed by the genotyping by sequence (GBS) data of 201 RILs from this population. A single marker was selected to represent each of the clusters and the linkage map constructed with these markers covers a total length of 1022.4 cM with an average interval of approximately 0.7 cM between loci. Three of the large-effect loci controlling KW (located on 2HL, 6HL, and 7HL, respectively) identified from the BA population were also detected in the MA population under different environments. The locus on 6HL was detected in each of the experiments conducted for both populations thus was selected for developing near isogenic lines (NILs). Apart from KW, the two isolines for each pair of the putative NILs obtained showed no significant difference for any of the morphological characteristics assessed. The average difference in KW between the isolines for all the NILs obtained was about 15% based on assessments under both glasshouse and field conditions. Taken advantage that high quality genome assemblies for both Morex and AWCS276 are available, we identified candidate genes underlying two of the three loci based on an orthologous analysis. The NILs developed and the candidate genes identified in this study should facilitate the cloning and functional analysis of genes regulating KW in barley.

1. Introduction

Barley (*Hordeum vulgare* L.) was one of the first domesticated grains and it is widely cultivated in temperate regions worldwide. The global production of barley in the 2017–2018 crop years was about 142 million metric tons. Barley is commonly used in beverage production, as animal feed, as well as edible grain in bread, cereal and soup [1]. Like other cereal crops, barley yield is determined by kernel weight (KW) and kernel number harvested per unit area. KW also affects both malt yield and feed quality in barley [2,3].

Similar to the situations in other crop species, KW in barley has been intensively investigated [4–7]. Results from QTL mapping studies indicate that the genetic control of KW is likely to be highly quantitative as several loci can routinely be detected from a single population [7]. Further, results from previous studies also show that strong interactions may exist between a targeted characteristic and other traits [8–11]. In regarding to KW in barley, it can be affected by grain size, endosperm hardness and grain density [5]. It can also be affected by inflorescence structure [7,12–14]. Due to the heterogeneity in genetic background in segregating populations routinely used for QTL mapping, lines containing a given allele do not always perform better than those containing the

* Corresponding authors.
E-mail addresses: liuyaxi@sicau.edu.cn (Y. Liu), chunji.liu@csiro.au (C. Liu).

[1] These authors contributed equally to this paper.

https://doi.org/10.1016/j.cj.2020.07.010

alternative allele thus average performances between lines with or without a target allele must be compared.

The existence of possible interactions between the characteristic under study and other traits mean that QTL detected in a mapping study can only be treated as putative and they need to be validated [15–17]. Near isogenic lines (NILs) are powerful genetic stocks that can be effectively used to validate and evaluate the effects of a given locus in different genetic backgrounds [11,15,16]. With the availability of linked markers, the method based on the heterogeneous inbred family (HIF) analysis has been widely used in generating NILs for various characteristics [11,15,18].

In the study reported here, we mapped loci controlling KW in barley using two populations of recombinant inbred lines (RILs). One of the populations was generated from two genotypes for which the whole genome assemblies were available, making it possible to identify candidate genes for each of the loci detected through sequence homology analysis. We also generated NILs targeting the most stable locus detected in both populations and used them to validate the effect of this loci in different genetic backgrounds.

2. Materials and methods

2.1. Plant materials

Two populations of RILs were used to detect QTL for KW in this study. AWCS276, a wild barley (*H. spontaneum* L.) was the shared parent between the two populations (Baudin/AWCS276 (BA) and Morex/AWCS276 (MA)). The linkage map for the BA population was reported in a previous study [19]. The MA population, containing 201 RILs, was generated from a cross between Morex and AWCS276 using the single-seed-descendent method with eight cycles of self-pollination based on the accelerated generation technique [20].

NILs were developed based on the method of HIF described earlier [15,21]. A single marker linked closely to the targeted locus was selected for developing the NILs. PCR reactions for the marker (Table S1) were performed in 25 μL mixtures containing 50 ng genomic DNA, 0.4 μmol L^{-1} of forward and reverse primer, 3 mmol L^{-1} $MgCl_2$, 1 mmol L^{-1} dNTPs, and 2.5 U *Taq* DNA polymerase. The PCR program included an initial denaturation at 94 °C for 5 min, followed by 35 cycles of 94 °C for 1 min, 60 °C for 1 min, 72 °C for 1 min, and a final extension at 72 °C for 10 min. The PCR products were mixed with Gel-Red stain and separated on 2.5% agarose gels. The marker was used to identify heterozygous plants from F_4 lines of the MA population. The identified plants were self-pollinated, and eight plants derived from each of them were used for the next round of selection. This process of identifying heterozygous individuals and self-pollination was repeated until F_9 generation. Two lines, one with the allele from Morex and the other with the allele from AWCS276, were then selected from each of the heterozygous individuals and treated as a putative NIL pair.

2.2. Field and glasshouse experiments

The BA population was evaluated in three consecutive seasons at the Wenjiang Research Station in Chengdu, China (30°36′N/103°41′E) in 2013 (WJ13), 2014 (WJ14), and 2015 (WJ15), respectively. Each of these experiments consisted of two replicates. Each replicate had ten plants grown in a single row (1.5 m in length), and distance between rows was 15 cm. Five plants from the middle of each row were harvested for measuring KW.

The MA population was assessed once in a field trial and twice in the glasshouses. The field trial was conducted at the CSIRO Gatton Research Station (27°33′S, 152°16′E) in June 2019. It consisted of two experiments, one with 4-week-vernalized seedlings (experiment GA19R1) and the other 6-week-vernalized seedlings (experiment GA19R2). Seeds were germinated in Petri dishes on three layers of filter paper saturated with water and then placed in a 6 °C cold room with constant lighting for vernalization. Each of the experiments contained two replicates, each with ten vernalized seedlings spaced (20 cm) planted in a single row. Row spacing was 25 cm. The two glasshouse trials were all conducted at the Queensland Bioscience Precinct (QBP) in Brisbane, one in December 2018 (experiment GH18) and the other in July 2019 (experiment GH19). Each of the glasshouse experiments consisted of three replicates. Three plants, each in a separate 2.0 L pot, were used in each of the replicates. A randomized complete block design was used for each of the experiments. Seeds were harvested from mature plants and dried in a 32 °C oven for ten days before being used for measuring KW. KW was based on the average of three replicated measurements of 100 randomly chosen kernels from each line used in these experiments.

The NILs were assessed in two experiments, one in a QBP glasshouse (experiment GH19) and the other at the CSIRO Gatton Research Station in June 2019 (experiment GA19). GH19 consisted of three replicates, each with five plants in different pots of 2.0 L. GA19 also consisted of three replicates, each with five plants growing in a single row. The plants within each row have a spacing of 20 cm, and the distance between rows was 25 cm. Kernel length (KL), kernel width (KWI), flag leaf width (FLW), flag leaf length (FLL), plant height (PH), spike exertion length (PEL), spike length (SL), tiller number (TN), and kernel number per spike (KNS) were measured. Measurements were taken from the main tillers of the middle three plants in each row and used for analysis.

2.3. Phenotypic data analysis

Significance test was conducted using the Student's *t*-test. Broad-sense heritability (H^2) for each trait was estimated as $H^2 = \sigma_g^2/(\sigma_g^2 + \sigma_{ge}^2/n + \sigma_e^2/nr)$, where σ_g^2 is the genetic variance, σ_{ge}^2 is the G × E variance, σ_e^2 is the error, n is the number of environments, and r is the number of replicates [22]. The σ_g^2, σ_{ge}^2, and σ_e^2 values were obtained based on the analysis of variance (ANOVA) using SAS 9.2 (SAS Institute Inc., Cary, NC, USA).

2.4. Genotyping by sequencing (GBS), SNP calling and linkage map construction

For the MA population, total genomic DNA was isolated from young leaves of all the RILs and the two parents using the CTAB method. The GBS data were generated by the Agriculture Victoria Research, Australia. Paired-end reads for each sample were error corrected and overlapped to produce single reads using the PEAR software with default parameters [23]. PEAR sequences from each sample were trimmed of the GBS probe sequences, concatenated and filtered to produce a unique list of allele sequences referred to as the population-specific allele-specific reference (psASR). Filtering applied included rejecting allele sequences having a frequency of occurrence of <4 in any sample or a minor allele frequency of <0.25 across the samples. The psASR sequences where aligned to the Morex genome assembly using the GYDLE software (Gydle Inc.). Alignment parameters included requirement for a minimum of 70% homology between the psASR sequence and reference genome across 95% of the psASR sequence. Exhaustive searching of the alignment space ensured that GYDLE returned the best alignment for each psASR sequence. Alignment information for each psASR sequence was used to identify alleles

originating from the same locus in order to facilitate codominant genotype calling. Those psASR sequences not aligned to the Morex genome assembly or had multiple hits with equal alignment score were classified as unmapped loci and treated as dominant markers for genotype calling. PEAR sequences for each sample were aligned to the psASR using GYDLE software using parameters that allowed up to 2 mismatches and guaranteed reporting of the best alignment. Those PEAR sequences having more than one hit to the psASR with the same alignment score were discarded. Genotype calling for mapped loci with multiple alleles was performed using a custom script that implements Bayesian probability statistics to call codominant genotypes based on the number of PEAR sequences aligned to each allele at a given locus. Genotype calling for mapped loci with only one allele (i.e. monomorphic loci) or unmapped loci was performed using a custom script that applies a minimum read coverage of 2 to call dominant genotypes. Raw SNP calls with minor allele frequency (MAF) of >0.05 in the biparental population were concatenated into a single variant call file (VCF) and used for linkage map construction. SNPs with more than 15% missing rate and those that were monomorphic between the two parents were eliminated. The retained SNPs were sorted into groups using MSTMap in the R/ASMap package with the default parameters [24].

2.5. QTL analysis and candidate genes prediction

QTL analysis was conducted using MAPQTL 6.0 [25]. A test of 1000 permutations was used to identify the LOD threshold, corresponding to a genome-wide false discovery rate of 5% ($P < 0.05$).

QTL detected in multiple environments were used to delineate the physical intervals. Tag sequences in GBS dataset were used to Blast against the Morex Assembly of RefSeq v2.0 [26] to get physical positions. Gene lists in the candidate regions were downloaded from the website ftp://ftp.ensemblgenomes.org/pub/release-44/plants/gff3/ hordeum_vulgare. At the same time, sequences of kernel-related genes from rice were collected and used to blast against Morex genes in the putative QTL intervals. Candidate genes detected in Morex were then used to Blast against the genome assembly of AWCS276 [27] to detect sequence differences between the two parents of the MA population.

3. Results

3.1. Phenotypic variation

Of the three parental genotypes for the two RIL populations used in this study, AWCS276 has the lowest KW (Table 1; Fig. 1). Transgressive segregation with normal distribution for KW was detected in each of the experiments conducted for both populations. Phenotypic variation coefficients ranged from 18.1% to 20.3% among the experiments for the BA population and ranged from 19.4% to 23.8% for the MA population (Tables 1, S2, and S3; Fig. S1). The broad-sense heritability was 0.681 and 0.684 for BA and MA, respectively, suggesting that genetic effects are the major determinant of the phenotypic variance on KW.

3.2. Genotyping and linkage map construction for the Morex/AWCS276 population

Approximately 173 million sequencing reads were obtained from 201 RILs of the MA population. The numbers of reads for each of these lines varied from 301,340 to 2178,836 with an average of nearly 0.85 million. A total of 21,598 SNPs with >25% minimum allele frequency (MAF) were identified from the GBS data. Following the discarding of those monomorphic and those missing more than 15% of the values, 5273 markers were retained and they formed 1454 clusters. A single marker was used to represent each of these clusters in linkage map construction. They formed seven linkage groups corresponding to each of the chromosomes. The genetic lengths of the linkage groups varied from 115.7 cM (4H) to 178.2 cM (2H). The total length of the linkage map is about 1022.4 cM with an average distance of 0.7 cM between markers (Tables S4 and S5). SNP distribution and characteristics of the linkage map are presented in detail in Tables S3 and S4. For the BA population, we used the genetic map reported in a previous study [19].

3.3. QTL analysis

Five QTL for KW were detected from the BA population from the experiments conducted over the three seasons. They were located on chromosomes 2H, 3H, 3H, 6H, and 7H, respectively. These QTL contributed between 10.5 (*KW-BA-2H*) and 30.3% (*KW-BA-3H.2*) of the phenotypic variance among the experiments. Only one of them, *KW-BA-6H*, was constantly detected in each of the experiments. LOD scores for this locus varied from 3.8 to 5.8 among the three experiments (Table 2; Fig. S2).

There were also five QTL for KW detected from the MA population. They located on 2H, 3H, 6H, 7H, and 7H, respectively. Two of them, including *KW-MA-2H* and *KW-MA-6H,* were detected in each of the four experiments. *KW-MA-3H* on chromosome arm 3HS was detected in two of the experiments. *KW-MA-7H.1* on chromosome arm 7HS and *KW-MA-7H.2* on 7HL were both detected in only one of the experiments (Table 2; Fig. 2).

Three of the QTL detected between the two populations likely shared similar physical locations: *KW-BA-2H* and *KW-MA-2H* located in the interval of 627–700 Mb on chromosome arm 2HL, *KW-BA-6H* and *KW-MA-6H* in the interval of 450–545 Mb on 6HL, and *KW-BA-7H* and *KW-MA-7H.2* in the interval of 593–637 Mb on 7HL. These were the three loci we focused on in further analyses in the study reported here and they were named as *KW-2H*, *KW-6H*, and *KW-7H*, respectively. The alleles for increased KW for all the three QTL were derived from the female parents (Baudin and Morex, respectively).

3.4. Development and assessment of NILs targeting the locus KW-6H

NILs targeting the locus *KW-6H* were obtained based on the approach of HIF analysis [32] with the use of one closely linked marker. Nine pairs of putative NILs were obtained. Significant difference in KW between the isolines for seven of these putative NILs was detected in both the assessments conducted. The differences between the isolines for each of the NIL pairs ranged from 9.7% to 23.3% with an average of 15.5% (Table 3). Consistent differences for any of the other characteristics assessed (including KL, KWI, FLW, PEL, SL, TN, FLL, PH, and KNS) were not detected between the isolines for any of the NIL pairs (Table 3; Fig. S3).

3.5. Prediction of candidate genes for KW-2H, KW-6H, and KW-7H

Sequences for 71 kernel-related genes from rice were identified. They were used to BLAST the assembled genomes of Morex and AWCS276. This analysis detected three orthologs (including *HORVU2Hr1G102800*, *HORVU2Hr1G096350*, and *HORVU2Hr1G099220*) in the *KW-2H* interval. These genes were orthologous to rice genes *SPL33*, *SMG2*, and *SPL29*, respectively. Sequence similarities of the three genes between rice and barley ranged from 75.1% to 85.9%. There were 2 SNPs (G → A) between Morex and AWCS276 in an exon of *HORVU2Hr1G102800*, and they were both synonymous mutations. Three SNPs in exons and a 260 bp InDel in the second intron were detected for *HOR-*

Table 1
Phenotypic performance and distribution of kernel weight among the three parents and two populations of recombinant inbred lines used in this study.

Population	Experiment	Parents		Recombinant inbred lines						
		FP	MP	Mean ± SD	Min	Max	Skew	Kurt	CV (%)	H^2
Baudin/AWCS276	WJ13	2.5 ± 0.1	4.3 ± 0.1	3.4 ± 0.7	1.1	5.2	−0.4	0.8	20.00	0.681
	WJ14	1.8 ± 0.1	3.0 ± 0.2	2.5 ± 0.5	1.4	3.7	0	−0.7	20.30	
	WJ15	2.8 ± 0.1	4.1 ± 0.2	3.4 ± 0.6	1.7	5.1	−0.1	−0.2	18.10	
Morex/AWCS276	GH18	2.2 ± 0.1	4.5 ± 0.1	3.8 ± 0.8	1.6	5.8	−0.1	0	22.00	0.684
	GH19	2.3 ± 0.3	4.8 ± 0.1	3.1 ± 0.7	1.5	4.9	0.1	−0.7	23.80	
	GA19R1	2.9 ± 0.1	4.1 ± 0.1	3.2 ± 0.6	1.7	5	0.2	−0.1	19.40	
	GA19R2	2.8 ± 0.2	3.7 ± 0.1	3.0 ± 0.6	1.5	4.5	0	−0.5	20.60	

FP, female parent; MP, male parent; SD, standard deviation; CV, coefficient of variation.

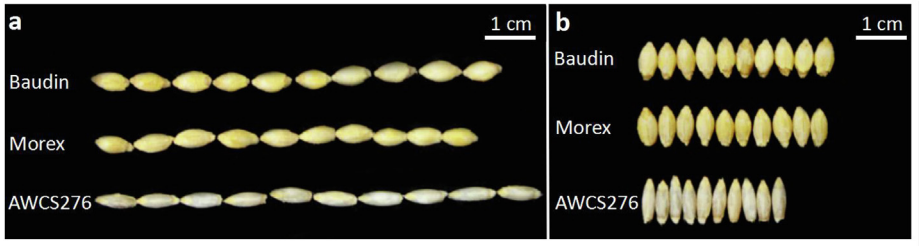

Fig. 1. Seed morphologies of the three parental lines for the two RIL populations used in this study.

Table 2
QTL for barley kernel weight detected in each of the two populations of recombinant lines in each of the environments.

QTL	Env.	Flanking marker		Linkage map interval (cM)	Physical map interval (Mb)	LOD	PVE (%)	References
		Left	Right					
Baudin/AWCS276								
KW-BA-2H	WJ13	M354	–	130.9	664	3.1	10.5	Xue et al. [28], Lex et al. [29], Wang et al. [30]
KW-BA-3H.1	WJ13	M809	M904	17.5–45.0	504–514	4.7	15.5	
KW-BA-3H.2	WJ14	M644	M713	15.8–41.2	611–650	10.2	30.3	Wang et al. [30]
KW-BA-6H	WJ13	M1437	M1445	61.1–67.4	498–507	3.9	12.9	Lex et al. [29], Walker et al. [5]
	WJ14	M1428	M1437	55.1–61.1	478–498	4.1	13.4	
	WJ15	M1416	M1479	48.3–75.3	450–510	5.8	19.7	
KW-BA-7H	WJ13	M1739	M1771	19.1–28.5	617–637	4.4	14.5	Wang et al. [30], Arifuzzaman et al. [31]
Morex/AWCS276								
KW-MA-2H	GH18	GBS_MST1275	GBS_MST1358	90.0–104.8	651–670	10.6	22.2	Xue et al. [28], Wang et al. [30]
	GH19	GBS_MST1270	GBS_MST1355	89.1–104.2	647–668	13.5	27.5	Lex et al. [29]
	GA19R1	GBS_MST1166	GBS_MST1375	80.7–109.8	636–687	17.2	33.9	
	GA19R2	GBS_MST810	GBS_MST1378	71.0–110.9	627–700	29.1	51.8	
KW-MA-3H	GH18	GBS_MST1533	GBS_MST1537	0–2.7	1–2	4.8	10.5	Wang et al. [30]
	GA19R1	GBS_MST1533	GBS_MST1536	0–1.4	1–2	3.1	7.1	
KW-MA-6H	GH18	GBS_MST4053	GBS_MST3973	65.4–85.6	506–545	5.0	10.9	Lex et al. [29], Walker et al. [5]
	GH19	GBS_MST4047	–	66.5	509	3.1	7.0	
	GA19R1	GBS_MST4371	GBS_MST3978	47.2–82.3	499–541	4.5	10.0	
	GA19R2	GBS_MST3978	GBS_MST4052	66.2–82.3	508–541	3.8	9.0	
KW-MA-7H.1	GH18	GBS_MST5183	GBS_MST5161	46.6–52.4	46–85	3.4	7.5	Wang et al. [30]
KW-MA-7H.2	GH18	GBS_MST4657	GBS_MST4611	104.1–113.7	593–613	4.0	8.8	Wang et al. [30], Arifuzzaman et al. [31]

VU2Hr1G096350 between the two parental genotypes. They included one C → G transversion and two T → C transitions. Of these three SNPs, only one was non-synonymous and it produces an amino acid residue substitution at position 763 (Proline → Alanine). Two SNPs (G → A) were found in an exon of *HORVU2Hr1G099220* between the two parents. The one at position 1411 produced an amino acid residue substitution (Alanine → Threonine). Only a single ortholog (*HORVU6Hr1G068370*) was identified in the *KW-6H* interval and it is homologous to the rice gene *OsGRF4* with a sequence similarity of 62.0%. There was only one SNP (G → T) in the last exon which leaded to an amino acid substitution (Aspartic acid → Tyrosine). There were also three insertion-deletions in the upstream region of the translation start codon between the two parental genotypes. No orthologs controlling KW in rice were detected in the interval harbouring the *KW-7H* locus (Table 4; Fig. 3).

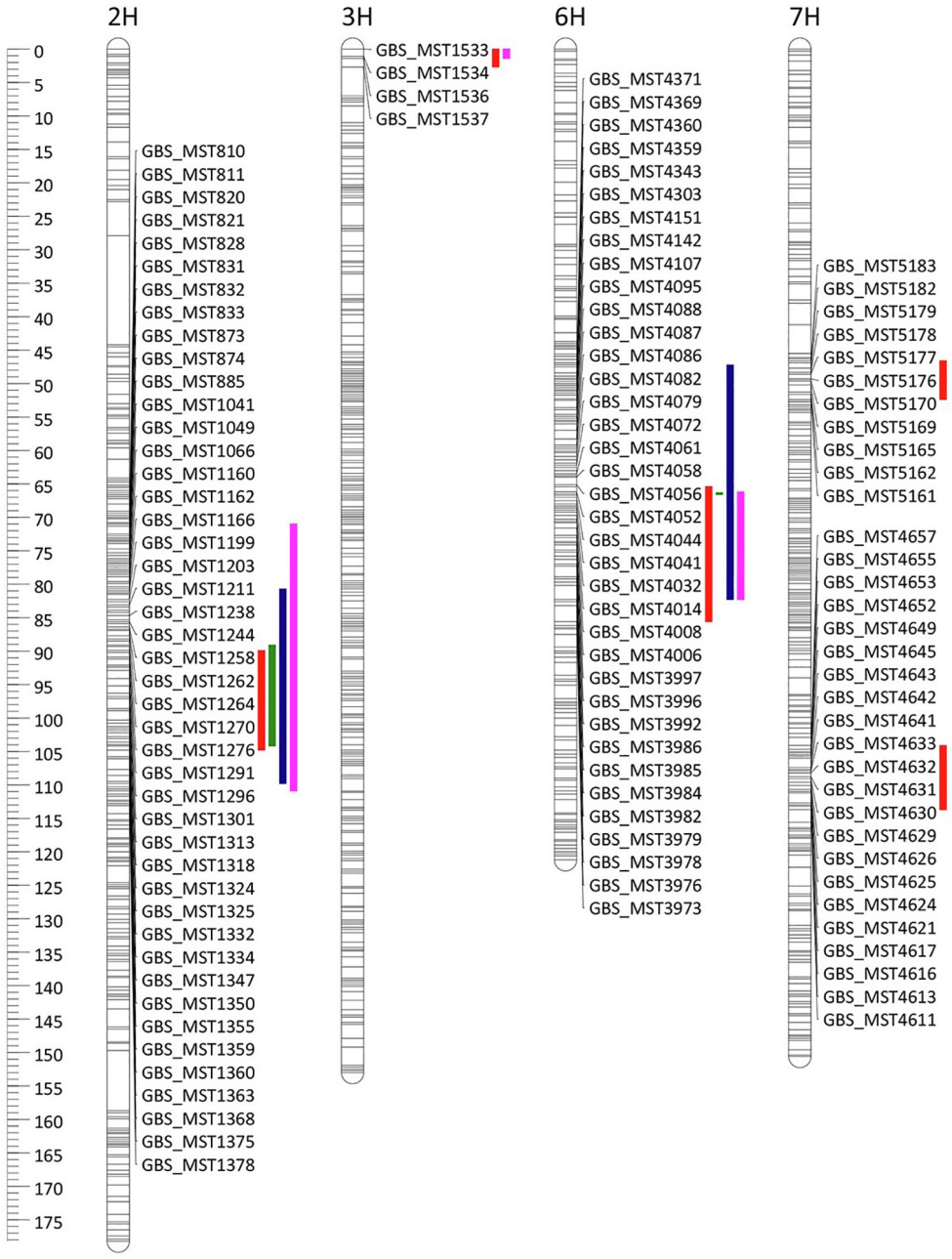

Fig. 2. Locations of QTL for kernel weight in the Morex/AWCS276 population. QTL regions are marked with bars in different colours. QTL detected in GH18 (red), in GH19 (green), in GA19R2 (blue), in GA19R1 (pink).

4. Discussion

KW is a key component affecting not only kernel yield but also malting quality in barley [2,3]. By assessing two populations of recombinant inbred lines with a shared parent, we investigated the genetic control of KW in the study reported here. One of these populations was developed as part of this study between two genotypes for which whole genome assemblies were available. GBS data for 201 of the RILs from the MA population was used in generating the linkage map which had an average interval of about 0.7 cM between loci. Three of the large-effect loci controlling KW, located on 2HL, 6HL, and 7HL, respectively, were detected from both populations. NILs targeting the locus on 6HL were generated based on the method of HIF with the use of a single marker. Apart from KW, significant differences for any other characteristics assessed were not detected between the two isolines for any of the NIL pairs obtained. Significant differences in KW between the isolines were detected for seven of the NIL pairs. The presence of the AWCS276 allele reduced the KW by an average of 15.3% based on results from the glasshouse assays and by an average of 15.5% based on results from the field assay.

It is of interesting to note that, although many loci conditioning KW have been reported, none of them seems to be more distally located than any of the three loci reported here. The locus on 2HL is in a similar region as those described by Xue et al. [28], Lex et al. [29], and Wang et al. [30]. Ren et al. [39] also reported a locus on this chromosome arm but it was about 130 Mb more proximally located. The locus on 6HL locates in a similar region as the ones reported by Walker et al. [5] and Lex et al. [29]. Arifuzzaman et al. [31] also reported a locus on this chromosome arm but

Table 3
Difference in various characteristics measured between the two isolines for each of the seven pairs of near isogenic lines targeting the kernel weight locus on chromosome arm 6HL.

Isoline	Trial	KW	KL	KWI	PH	PEL	FLL	FLW	SL	TN	KNS
NIL1											
Morex	GH19	4.6	10	3.4	93.7	12.1	14.4	0.6	8.3	25.5	16.5
A276		4	10.2	3.3	96	11.9	15.6	0.7	8.6	21.5	17
Diff. (%)		15.12***	−2.19	2.46	−2.45	2.11	−8.01	−21.43	−3.49	18.60	−2.94
Morex	GA19	4.4	10.1	3.3	65.5	0	10.5	0.7	5.4	49	
A276		4	9.9	3.2	73.5	0	8.9	0.5	4.5	42.7	
Diff. (%)		9.7***	2.10	2.90	−11	0.00	18.60	31.49	20.22	14.84	
NIL2											
Morex	GH19	5.4	10.7	3.5	94.6	16.2	16.1	0.5	7.8	19	17
A276		4.5	10.3	3.4	93.6	18.2	12.8	0.5	6.8	18	16
Diff. (%)		18.65*	3.87	1.46	1.07	−11	25.78**	0.00	14.71**	5.56	6.25
Morex	GA19	3.85	9.64	3.14	55.6	0	11.67	0.7	8.66	44	
A276		3.41	9.89	3.15	60.4	0	11.07	0.64	7.87	42	
Diff. (%)		12.9**	−2.53	−0.32	−7.95	0.00	5.42	9.38	9.99	4.76	
NIL3											
Morex	GH19	5.9	9.4	3.6	92.6	13.5	13.5	0.7	7.4	19	18
A276		5.2	9.3	3.6	86.1	14.5	13.4	0.8	7.6	11	18
Diff. (%)		12.85**	1.19	−0.14	7.55*	−6.7	0.75	−12.5	−2.63	72.73***	0.00
Morex	GA19	4.2	9.3	3.4	71.8	0	10.6	0.8	10.3	46	
A276		3.4	9	3.1	44.6	0	9.8	0.8	8	27	
Diff. (%)		20.84***	2.70	9.04**	61**	0.00	7.39	5.63	28.12**	70.37***	
NIL4											
Morex	GH19	5.9	9.3	3.6	93.8	12.4	19.9	0.7	7.8	16	19
A276		5.3	9.3	3.6	91.4	12.5	16.4	0.7	7.3	15.5	19
Diff. (%)		12.11**	−0.05	−0.62	2.68	−0.4	21.34**	0.00	7.59	3.23	0.00
Morex	GA19	4.1	9	3.2	56.9	0	9.4	0.5	8.6	42	
A276		3.6	9.1	3.1	56.2	0	9.3	0.7	8.5	46.3	
Diff. (%)		14.55***	−1.28	4.32	11.90	0.00	0.50	−28.97*	1.58	−9.353	
NIL5											
Morex	GH19	5.6	10.3	3.5	96.1	12.5	7.9	0.4	8.6	13	18
A276		4.5	10.1	3.4	97.9	18.3	10.6	0.5	7.8	18.7	18.3
Diff. (%)		23.34**	0.51	0.97	−1.81	−31.57**	−25.47**	−14.29	10.73*	−30.36**	−1.82
NIL6											
Morex	GH19	5.9	9.2	3.6	90.8	11.1	9.9	0.6	7.1	11	18
A276		5.1	9.3	3.6	92.7	16.1	13.9	0.7	7.2	12.5	19.5
Diff. (%)		14.71**	−1.55	−0.69	−2.05	−30.84**	−28.78**	−14.29	−1.39	−12**	−7.69
NIL7											
Morex	GH19	5.4	9.2	3.6	87.2	11.6	15.8	0.8	7.6	14.3	17.3
A276		4.9	9.1	3.5	82.1	11.8	14	0.8	7.7	18	16
Diff. (%)		10.12*	1.11	4.05	6.21	−1.69	12.5*	2.56	−0.97	−20.83**	7.81

Morex and A276, allele from Morex and AWCS276, respectively; Diff., difference; KW, kernel weight; KL, kernel length; KWI, kernel width, PH, plant height; PEL, spike exertion length; FLL, flag leaf length; FLW, flag leaf width; SL, spike length; TN, tiller number; KNS, kernel number per spike. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

Table 4
High confidence genes in the targeted intervals of barley and their orthologs in rice.

QTL	Barley orthologs	Physical position (bp)	Rice gene	Identity	SNPs[#]	Amino acids[#]	References
KW-2H	*HORVU2Hr1G096350*	chr2H: 674144117–674153272	*SMG2*	0.859	C/G (763), T/C (1695), T/C (1749)	P–A (255)	Guo et al. [33], Xu et al. [34]
	HORVU2Hr1G099220	chr2H: 687144366–687149157	*SPL29*	0.804	G/A (462), G/A (1411)	A–T (471)	Wang et al. [35]
	HORVU2Hr1G102800	chr2H: 698219940–698225810	*SPL33*	0.751	G/A (612), G/A (765)	None	Wang et al. [36]
KW-6H	*HORVU6Hr1G068370*	chr6H: 473135773–473139363	*OsGRF4*	0.62	G/T (1111)	D–Y (371)	Sun et al. [37], Duan et al [38]

[#] The numbers in brackets represent the positions of differences in nucleotide or amino acid sequences between Morex and AWCS276 relative to initiation codons.

it located about 100 Mb more proximally located than the one described here. The one on chromosome arm 7HL locates in a similar region as those reported by Arifuzzaman et al. [31] and Wang et al. [30]. Ren et al. [39] also reported a locus on this chromosome arm but it located about 105 Mb proximally compared with the one described here. It is well known that recombination rates can vary among populations [40–42] but they do not explain whether the more distal locations for all the three loci reported here are simply a coincidence.

The initial method of HIF for developing NILs was based on the use of flanking markers. If both the flanking markers are closely linked with their targeted locus, the rate of failure for this approach should be rare. However, linkage drag can be a significant issue when flanking markers derived from QTL mapping are used. This is because that the minimum difference between the two isolines for a given NIL pair would be the full length of the chromosomal fragment flanked by the markers. The fragment can be large if the markers are not tightly linked with the targeted locus, which is often the case when only QTL mapping results are available in selecting the markers. To obtain NILs with reduced sizes of the chromosome segments differentiating the isolines, a single marker linked with a targeted locus can be used [15,18]. However,

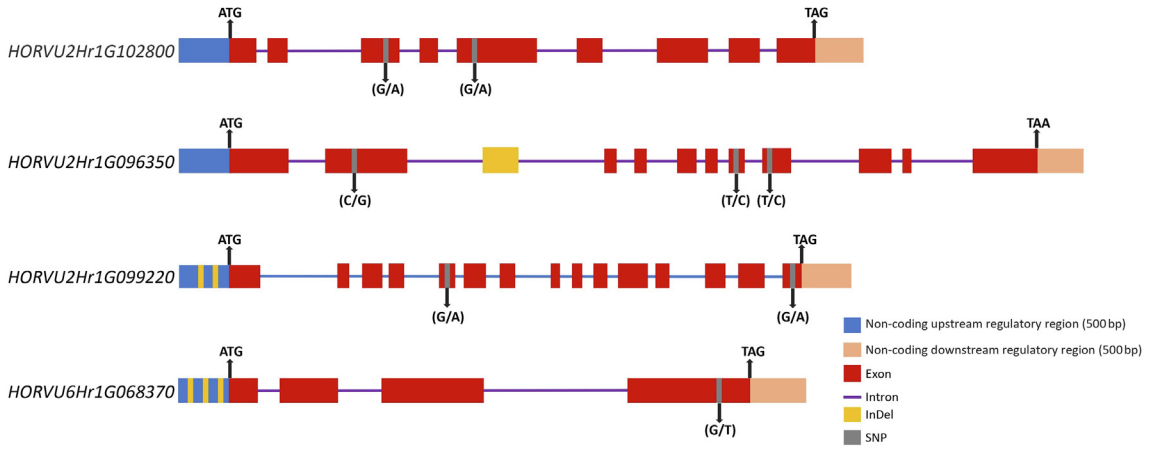

Fig. 3. Gene structures and genome sequence alignments of candidate genes for KW-2H and KW-6H between Morex and AWCS276.

recombination between a marker and its targeted locus could occur thus some of the NILs obtained using a single linked marker could be false which do not segregate at the targeted locus. This is the likely reason why, similar to the results obtained in previous studies [18,43], two of the putative NIL pairs generated in this study failed to show significant difference between their isolines for KW.

Sequence comparisons between Morex and AWCS276 for genes with known effects on KW in rice identified candidate genes underlying KW for two of the three loci reported in this study. Assessing the functions of these candidate genes could lead to rapid isolation of genes underlying these loci. These results demonstrate the huge advantage of genetic mapping based on populations for which the genome assemblies of their parental genotypes are available. With the rapid advancement in sequencing capacity, it should not take long before this approach can be routinely used for crop species with large and complex genomes thus dramatically enhancing our ability to identify and clone genes from such species.

Declaration of competing interest

Authors declare that there are no conflicts of interest.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (31771794), the National Key Research and Development Program of China (2017YFD0100900 and 2016YFD0101004), the Outstanding Youth Foundation of the Department of Science and Technology of Sichuan Province (2016JQ0040), and the International Science & Technology Cooperation Program of the Bureau of Science and Technology of Chengdu China (2015DFA306002015-GH03-00008-HZ). The authors are also grateful to Drs. Matt Hayden and Kerrie L. Forrest at Agriculture Victoria Research for helping in generating the GBS data on the mapping population. Hong Zhou, Wei Luo, and Shang Gao are grateful to the Sichuan Agricultural University and the China Scholarship Council for funding their visit to CSIRO.

Appendix A. Supplementary data

Supplementary data for this article can be found online at https://doi.org/10.1016/j.cj.2020.07.010.

References

[1] A.C. Newton, A.J. Flavell, T.S. George, P. Leat, B. Mullholland, L. Ramsay, C. Revoredo-Giha, J. Russell, B.J. Steffenson, J.S. Swanston, Crops that feed the world 4. Barley: a resilient crop? Strengths and weaknesses in the context of food security, Food Secur. 3 (2011) 141.
[2] Y. Gan, E.H. Stobbe, Seedling vigor and grain yield of 'Roblin' wheat affected by seed size, Agron J. 88 (1996) 456–460.
[3] C.K. Walker, J.F. Panozzo, Development of a small scale method to determine volume and density of individual barley kernels, and the relationship between grain density and endosperm hardness, J. Cereal Sci. 54 (2011) 311–316.
[4] X. Gong, R. Wheeler, W.D. Bovill, G.K. McDonald, QTL mapping of grain yield and phosphorus efficiency in barley in a Mediterranean-like environment, Theor. Appl. Genet. 129 (2016) 1657–1672.
[5] C.K. Walker, R. Ford, M. Muñoz-Amatriaín, J.F. Panozzo, The detection of QTLs in barley associated with endosperm hardness, grain density, grain size and malting quality using rapid phenotyping tools, Theor. Appl. Genet. 126 (2013) 2533–2551.
[6] N.A. Tinker, D.E. Mather, B.G. Rossnagel, K.J. Kasha, A. Kleinhofs, P.M. Hayes, D. E. Falk, T. Ferguson, L.P. Shugar, W.G. Legge, Regions of the genome that affect agronomic performance in two-row barley, Crop Sci. 36 (1996) 1053–1062.
[7] D. Sun, Q. Wang, G. Sun, X. Ren, B. Du, Y. Cheng, Y. Wang, C. Li, Dissecting the genetic basis of grain size and weight in barley (*Hordeum vulgare* L.) by QTL and comparative genetic analyses, Front. Plant Sci. 10 (2019) 469.
[8] W. Yan, H. Li, S. Cai, H. Ma, G. Rebetzke, C. Liu, Effects of plant height on type I and type II resistance to Fusarium head blight in wheat, Plant Pathol. 60 (2011) 506–512.
[9] Y. Liu, X. Yang, J. Ma, Y. Wei, Y. Zheng, H. Ma, J. Yao, G. Yan, Y.G. Wang, J. Manners, Plant height affects Fusarium crown rot severity in wheat, Phytopathology 100 (2010) 1276–1281.
[10] H.B. Li, M.X. Zhou, C.J. Liu, A major QTL conferring crown rot resistance in barley and its association with plant height, Theor. Appl. Genet. 118 (2009) 903–910.
[11] G. Chen, W. Yan, Y. Liu, Y. Wei, M. Zhou, Y.L. Zheng, J.M. Manners, C. Liu, The non-gibberellic acid-responsive semi-dwarfing gene *uzu* affects Fusarium crown rot resistance in barley, BMC Plant Biol. 14 (2014) 22.
[12] L. Ramsay, J. Comadran, A. Druka, D.F. Marshall, R. Waugh, *INTERMEDIUM-C*, a modifier of lateral spikelet fertility in barley, is an ortholog of the maize domestication gene *TEOSINTE BRANCHED 1*, Nat. Genet. 43 (2011) 169–172.
[13] H. Bull, M.C. Casao, M. Zwirek, A.J. Flavell, W.T. Thomas, W. Guo, R. Zhang, P. Rapazote-Flores, S. Kyriakidis, J. Russell, Barley *SIX-ROWED SPIKE3* encodes a putative Jumonji C-type H3K9me2/me3 demethylase that represses lateral spikelet fertility, Nat. Commun. 8 (2017) 936.
[14] H.M. Youssef, K. Eggert, R. Koppolu, A.M. Alqudah, N. Poursarebani, A. Fazeli, S. Sakuma, A. Tagiri, T. Rutten, G. Govind, *VRS2* regulates hormone-mediated inflorescence patterning in barley, Nat. Genet. 49 (2017) 157.
[15] A. Habib, S. Shabala, L. Shabala, M. Zhou, C. Liu, Near-isogenic lines developed for a major QTL on chromosome arm 4HL conferring Fusarium crown rot resistance in barley, Euphytica 209 (2016) 555–563.
[16] S. Gao, Z. Zheng, J. Powell, A. Habib, C. Liu, Validation and delineation of a locus conferring Fusarium crown rot resistance on 1HL in barley by analysing transcriptomes from multiple pairs of near isogenic lines, BMC Genomics 20 (2019) 650.
[17] Y. Jiang, A. Habib, Z. Zheng, M. Zhou, Y. Wei, Development of tightly linked markers and identification of candidate genes for Fusarium crown rot resistance in barley by exploiting a near-isogenic line-derived population, Theor. Appl. Genet. 132 (2019) 217–225.
[18] J. Ma, G.J. Yan, C.J. Liu, Development of near-isogenic lines for a major QTL on 3BL conferring Fusarium crown rot resistance in hexaploid wheat, Euphytica 183 (2012) 147–152.

[19] H. Zhou, S. Liu, Y. Liu, Y. Liu, J. You, M. Deng, J. Ma, G. Chen, Y. Wei, C. Liu, Mapping and validation of major quantitative trait loci for kernel length in wild barley (*Hordeum vulgare* ssp. *spontaneum*), BMC Genet. 17 (2016) 130.
[20] Z. Zheng, H.B. Wang, G.D. Chen, G.J. Yan, C.J. Liu, A procedure allowing up to eight generations of wheat and nine generations of barley per annum, Euphytica 191 (2013) 311–316.
[21] G. Chen, Y. Liu, J. Ma, Z. Zhi, Y. Wei, M.I.C. Lynne, Y.L. Zheng, C. Liu, M. Xu, A novel and major quantitative trait locus for Fusarium crown rot resistance in a genotype of wild barley (*Hordeum spontaneum* L.), PLoS ONE 8 (2013) e58040.
[22] S.E. Smith, R.O. Kuehl, I.M. Ray, R. Hui, D. Soleri, Evaluation of simple methods for estimating broad-sense heritability in stands of randomly planted genotypes, Crop Sci. 38 (1998) 1125–1129.
[23] J. Zhang, K. Kobert, T. Flouri, A. Stamatakis, PEAR: a fast and accurate Illumina Paired-End reAd mergeR, Bioinformatics 30 (2014) 614–620.
[24] Y. Wu, P.R. Bhat, T.J. Close, S. Lonardi, Efficient and accurate construction of genetic linkage maps from the minimum spanning tree of a graph, PLoS Genet. 4 (2008) e1000212.
[25] J.W. van Ooijen, MapQTL 6, Software for the mapping of quantitative trait loci in experimental populations of diploid species, Kyazma BV, Wageningen, Netherlands, 2009.
[26] M. Mascher, H. Gundlach, A. Himmelbach, S. Beier, S.O. Twardziok, T. Wicker, V. Radchuk, C. Dockter, P.E. Hedley, J. Russell, A chromosome conformation capture ordered sequence of the barley genome, Nature 544 (2017) 427.
[27] M. Liu, Y. Li, Y. Ma, Q. Zhao, J. Stiller, Q.i. Feng, Q. Tian, D. Liu, B. Han, C. Liu, The draft genome of a wild barley genotype reveals its enrichment in genes related to biotic and abiotic stresses compared to cultivated barley, Plant Biotechnol. J. 18 (2020) 443–456.
[28] D.W. Xue, M.X. Zhou, X.Q. Zhang, S. Chen, K. Wei, F.R. Zeng, Y. Mao, F.B. Wu, G.P. Zhang, Identification of QTLs for yield and yield components of barley under different growth conditions, J. Zhejiang Univ. Sci. B 11 (2010) 169–176.
[29] J. Lex, J. Ahlemeyer, W. Friedt, F. Ordon, Genome-wide association studies of agronomic and quality traits in a set of German winter barley (*Hordeum vulgare* L.) cultivars using Diversity Arrays Technology (DArT), J. Appl. Genet. 55 (2014) 295–305.
[30] Q. Wang, G. Sun, X. Ren, B. Du, Y. Cheng, Y. Wang, C. Li, D. Sun, Dissecting the genetic basis of grain size and weight in barley (*Hordeum vulgare* L.) by QTL and comparative genetic analyses, Front. Plant Sci. 10 (2019) 469.
[31] M.d. Arifuzzaman, M.A. Sayed, S. Muzammil, K. Pillen, H. Schumann, A.A. Naz, J. Léon, Detection and validation of novel QTL for shoot and root traits in barley (*Hordeum vulgare* L.), Mol. Breed. 34 (2014) 1373–1387.
[32] M. Tuinstra, G. Ejeta, P. Goldsbrough, Heterogeneous inbred family (HIF) analysis: a method for developing near-isogenic lines that differ at quantitative trait loci, Theor. Appl. Genet. 95 (1997) 1005–1011.
[33] T. Guo, K.E. Chen, N.Q. Dong, C.L. Shi, W.W. Ye, J.P. Gao, J.X. Shan, H.X. Lin, *GRAIN SIZE AND NUMBER1* negatively regulates the *OsMKKK10-OsMKK4-OsMPK6* cascade to coordinate the trade-off between grain number per panicle and grain size in rice, Plant Cell 30 (2018) 871–888.
[34] R. Xu, P. Duan, H. Yu, Z. Zhou, B. Zhang, R. Wang, J. Li, G. Zhang, S. Zhuang, J. Lyu, Control of grain size and weight by the OsMKKK10-OsMKK4-OsMAPK6 signaling pathway in rice, Mol. Plant 11 (2018) 860–873.
[35] Z. Wang, Y. Wang, X. Hong, D. Hu, C. Liu, J. Yang, Y. Li, Y. Huang, Y. Feng, H. Gong, Functional inactivation of UDP-N-acetylglucosamine pyrophosphorylase 1 (UAP1) induces early leaf senescence and defence responses in rice, J. Exp. Bot. 66 (2014) 973–987.
[36] S. Wang, C. Lei, J. Wang, J. Ma, S. Tang, C. Wang, K. Zhao, P. Tian, H. Zhang, C. Qi, SPL33, encoding an eEF1A-like protein, negatively regulates cell death and defense responses in rice, J. Exp. Bot. 68 (2017) 899–913.
[37] P. Sun, W. Zhang, Y. Wang, Q. He, F.u. Shu, H. Liu, J. Wang, J. Wang, L. Yuan, H. Deng, OsGRF4 controls grain shape, panicle length and seed shattering in rice, J. Integr. Plant Biol. 58 (2016) 836–847.
[38] P. Duan, S. Ni, J. Wang, B. Zhang, R. Xu, Y. Wang, H. Chen, X. Zhu, Y. Li, Regulation of *OsGRF4* by *OsmiR396* controls grain size and yield in rice, Nat. Plants 2 (2015) 1–5.
[39] X. Ren, D. Sun, G. Sun, C. Li, W. Dong, Molecular detection of QTL for agronomic and quality traits in a doubled haploid barley population, Aust. J. Crop Sci. 7 (2013) 878–886.
[40] M.C. de Vicente, S.D. Tanksley, Genome-wide reduction in recombination of backcross progeny derived from male versus female gametes in an interspecific cross of tomato, Theor. Appl. Genet. 83 (1991) 173–178.
[41] C.S. Busso, C.J. Liu, C.T. Hash, J.R. Witcombe, K.M. Devos, J.M.J. de Wet, M.D. Gale, Analysis of recombination rate in female and male gametogenesis in pearl millet (*Pennisetum glaucum*) using RFLP markers, Theor. Appl. Genetics 90 (1995) 242–246.
[42] C.J. Liu, K.M. Devos, J.R. Witcombe, T.S. Pittaway, M.D. Gale, The effect of genome and sex on recombination rates in Pennisetum species, Theor. Appl. Genet. 93 (1996) 902–908.
[43] A. Habib, J.J. Powell, J. Stiller, M. Liu, S. Shabala, M. Zhou, D.M. Gardiner, C. Liu, A multiple near isogenic line (multi-NIL) RNA-seq approach to identify candidate genes underpinning QTL, Theor. Appl. Genet. 131 (2018) 613–624.